MERCATUS
RESEARCH

AN INTRODUCTION TO US MONETARY POLICY

Steven Horwitz

MERCATUS CENTER
George Mason University
Bridging the gap between academic ideas and real-world problems

ABOUT THE MERCATUS CENTER AT GEORGE MASON UNIVERSITY

The Mercatus Center at George Mason University is the world's premier university source for market-oriented ideas—bridging the gap between academic ideas and real-world problems.

A university-based research center, Mercatus advances knowledge about how markets work to improve people's lives by training graduate students, conducting research, and applying economics to offer solutions to society's most pressing problems.

Our mission is to generate knowledge and understanding of the institutions that affect the freedom to prosper and to find sustainable solutions that overcome the barriers preventing individuals from living free, prosperous, and peaceful lives.

Founded in 1980, the Mercatus Center is located on George Mason University's Arlington campus.

www.mercatus.org

Mercatus Center
George Mason University
3351 Fairfax Drive, 4th Floor
Arlington, VA 22201-4433
(703) 993-4930
mercatus.org

Release date: April 2, 2013

Corrected version

ABOUT THE AUTHOR

Steven Horwitz is an affiliated senior scholar at the Mercatus Center at George Mason University, as well as the Charles A. Dana professor of economics at St. Lawrence University in Canton, New York. He completed his MA and PhD in economics at George Mason University and received his AB in economics and philosophy from the University of Michigan. He is the author of two books, *Microfoundations and Macroeconomics: An Austrian Perspective* and *Monetary Evolution, Free Banking, and Economic Order*, and he has written extensively on Austrian economics, Hayekian political economy, monetary theory and history, and macroeconomics. The author of numerous op-eds, Horwitz is a frequent guest on TV and radio programs, speaking particularly on the Great Recession and monetary policy. A member of the Mont Pelerin Society, Horwitz has spoken to professional, student, policy-maker, and general audiences throughout the United States, Canada, Europe, and South America.

ABSTRACT

This study examines the history and operation of the Federal Reserve System ("the Fed"). It explores the Fed's origins in American economic history and emphasizes the political compromises that produced it. It seeks to provide an accessible explanation of how the Fed attempts to change the money supply and of the structural challenges it faces as it attempts to get the money supply correct. The paper uses the framework thereby developed to examine recent monetary policy, including quantitative easing. Inflation and deflation result when the Fed creates too much or too little money, and the study discusses the causes and costs of both in detail. The paper concludes with an examination of alternatives to central banking, including the gold standard and a system of competition in money production known as free banking.

JEL codes: E5, N1, N2

Keywords: Federal Reserve, monetary policy, quantitative easing, American economic history, financial institutions, inflation, deflation, gold standard, free banking

Many people find economics complicated to the point of being impenetrable. Of all the subject areas within economics, monetary theory and policy may be the most challenging for the general public. Money itself is a bit puzzling. After all, how did it happen that we give other people real goods and services and are willing to accept little pieces of paper or metal, or even bits and bytes on a computer, in return? When we expand that set of issues to include the workings of banks and other financial institutions, matters become even more abstract and complex. Few people understand how banks shift funds around and buy and sell sophisticated financial instruments. Adding monetary policy and the Federal Reserve System makes matters more complicated by an order of magnitude. The average citizen does not understand how the Fed makes its decisions, what it is trying to accomplish, or how it executes its plans.

In the text that follows, I hope to demystify many of these issues in monetary economics and monetary policy. Understanding what the Federal Reserve System ("the Fed") does today and how it attempts to achieve macroeconomic policy goals requires some knowledge of where the Fed came from, what tools it has at its disposal, and what various objectives it might be trying to accomplish. I will explore the history of monetary institutions in the United States to put the Fed's role in context and then look at the tools it can use to affect the money supply and, in turn, the macroeconomy. There are a variety of indicators that monetary policy makers might choose to affect, and which ones they target depends on their beliefs about the state of the economy and, more importantly, the particular macroeconomic theory they subscribe to. I will critically assess several of those frameworks and offer a superior alternative.

Monetary policy is important to understand, and to understand correctly, because mistakes can cause enormous problems. On the one hand, not creating enough money can lead to the kind of massive deflation we saw in the early 1930s that turned a severe recession into the Great Depression. On the other hand, creating too much money can lead to serious inflation like the United States saw in the 1970s and early 1980s and to the far worse inflation that other countries have experienced in the last few decades. Those high rates of inflation damage not just the macroeconomy, but also many of the households it comprises. Finally, I will

consider some alternatives to the Federal Reserve System and evaluate their potential for providing a stable monetary framework for sustainable economic growth by avoiding both deflation and inflation.

THE ORIGINS OF THE FEDERAL RESERVE AND THE MODERN US MONETARY SYSTEM

The history of American monetary institutions is one in which the role of government, particularly the federal government, has grown substantially, especially since the turn of the 20th century. The most common explanations for that growth follow a narrative similar to those that attempt to explain the origins of government growth in other areas of the economy. That narrative starts with the claim that there was a period of laissez-faire during which producers, in this case banks, took advantage of the lack of regulations to harm their customers and cause market-wide inefficiencies. These supposed "market failures" led to calls for government action, and defenders of government intervention claim that the subsequent regulations and new institutions solved the problems the free market created. Unfortunately, this narrative, most frequently told about the Progressive Era (roughly 1880–1914), is mistaken. It is particularly mistaken in the case of monetary institutions, since the problems that led to federal involvement, and to the creation of the Federal Reserve System in particular, were due to the combination of bad government regulations and the way in which bigger government served the self-interest of both bankers and politicians. I offer an alternative reading of this history.

The first problem with this narrative as applied to the United States is that there has never been a time when banking was unregulated. At the time of the nation's founding, banking and finance were regulated at the state level, and a few states prohibited banking altogether. The ones that allowed it regulated it heavily. Getting a bank charter in those states required legislative approval in each case. Charters became heavily politicized, and what we would today call "cronyism" was rampant. The result was a rather unsatisfactory banking system: banks were too few in number, their assets were insufficiently diversified, and they were often run by people who knew little about the business but who did know the right people. These problems were not the result of too much "laissez-faire."

In the 1830s, a number of states, starting with New York and Michigan, addressed these problems. They depoliticized getting a charter by setting up a clearly defined process that would give a charter to all banks that met the conditions. These so-called free banking laws were some of the few moves toward a less regulated banking system in the 19th century. However, even with this new chartering process, banks remained heavily regulated.

The three main regulations of the "Free Banking Era" from 1837 to 1863 were (1) limits on banks' ability to operate branches, (2) minimum reserve requirements, and (3) requirements that banks that produced currency buy up certain bonds or

other financial assets as collateral.[1] The first and third of these regulations were particularly problematic. During the Free Banking Era branch banking was largely prohibited. Smaller banks successfully convinced politicians that allowing branch banking would enable larger, urban banks to run them out of business, regardless of the potential benefits for consumers. The problem with small banks is that they tend to be very underdiversified, with most of their loans going to the specific industries in their limited geographic areas. This situation makes small banks especially prone to fail in the face of an economic downturn. Banks with large branching networks can diversify their loans across industries and geographic areas and move funds from areas doing well to areas that need help. Longstanding and misplaced fears of branching leading to some sort of money monopoly, particularly one controlled by New York City banks, kept the US banking system in a state of perpetual underdiversification, which in turn explains why US recessions and depressions have produced an inordinate number of bank failures compared to other countries.[2]

The branching restrictions also prevented the development of a truly integrated nationwide banking system. With banks limited to individual states, and sometimes individual offices, banks had difficulty developing the working relationships with banks in other states that are necessary for interstate commerce. For consumers, branching restrictions meant having to switch currencies and determine the value of each when traveling to a different state. Both banks and their customers often required the specialized services of note reporters for information about the exchange rates between different banks' notes. The costliness of engaging in banking across state lines made nationwide integration difficult.

The bond collateral requirements were also a problem. During the Free Banking Era, these requirements often served as a form of crony capitalism as some states required that banks buy the bonds of railroads and other nominally private enterprises instead of or in addition to government bonds to serve as collateral. Another problem was that the required bonds were sometimes found to be worthless, which contributed to the periodic currency panics that continued throughout the century.

The frustrations caused by the lack of a nationally integrated banking system led to calls for a federal solution. The start of the Civil War gave Congress an opportunity to both get involved with solving that problem and help finance the war effort. In 1863, Congress authorized the federal government to offer federal charters to interested banks. The new "national banks" would be regulated at the federal level, though they would still not be able to operate across state lines. The federal regulation would include setting reserve requirements and, crucially, a new set of bond

1. Those regulations indicate why the Free Banking Era should not be confused with the free banking system (to be discussed in the final section), which involves a radical deregulation of banking and finance as well as elimination of the central bank.
2. More information on the institutional history of the US monetary system can be found in Richard H. Timberlake, *Monetary Policy in the United States* (Chicago: University of Chicago Press, 1993).

collateral requirements. The law also required that national banks accept each other's currency at par, which was not only important for reducing the costs of valuing other banks' currency, but also prevented the discounting of notes from banks that were overissuing and thereby reducing the value of their notes. Federal politicians realized that requiring federally chartered banks to purchase federal government bonds as collateral for currency issue would provide a source of financing for the North's war effort. In fact, the use of banking regulation and central banks as a way to finance government expenditures, particularly for war, explains the origins of numerous central banks and other government interventions in banking around the world.

Even as the new national banking system shifted regulation to the federal government in an effort to integrate the banking system, it kept the same problematic regulations of the state-chartered system.[3] The result was a series of ever-worsening banking panics, culminating in severe panics in 1893 and 1907.[4] These panics were characterized by increases in the demand for the currency issued by the national banks, normally around harvest season when more currency was needed to buy and sell crops. The bond collateral requirements made issuing more currency expensive and time-consuming, as banks had to pay for the necessary bonds, often at high prices, and wait several weeks for the currency, which was physically printed in Washington, to be delivered. One reason the bonds became expensive was that the federal government began to retire debt after the Civil War, reducing the quantity of bonds available to serve as collateral and causing their prices to rise. By the time this whole currency issuing process was complete, customers would have long turned to currency substitutes or simply asked for gold instead of paper money. The demand for gold reserves caused the banks to draw down their reserves at other banks, and this ripple effect caused numerous banks to call in loans or borrow gold. The result in both 1893 and 1907 was recession. Critics, such as Paul Krugman, frequently argue that the problems of the pre-Fed US banking system reflect a failure of "laissez-faire."[5] As should be clear, this view is mistaken: even though there was no central bank, the federal government (as well as the states) nonetheless played a significant role in regulating the banking system. Both observers at the time and modern economic historians have pointed to these regulations, especially the limitations on branching and the bond collateral requirements, as the primary causes of

3. The state-chartered banks did not disappear. The creation of the Fed effectively ended the importance of the tax on state bank notes, leaving a more level playing field between a state charter and a federal charter. States continue to charter banks in the present day.
4. On these panics and the government's role in causing them, see Steven Horwitz, "Competitive Currencies, Legal Restrictions and the Origins of the Fed: Some Evidence from the Panic of 1907," *Southern Economic Journal* 56, no. 4 (January 1990): 639–49.
5. Paul Krugman, "Why We Regulate," *New York Times*, May 13, 2012, http://www.nytimes.com/2012/05/14/opinion/krugman-why-we-regulate.html?_r=2&.

the recurring panics and associated recessions. The problems with the US banking system were due to ill-conceived regulations, not laissez-faire.

The critics who see matters otherwise also ignore the nearby success story of Canada, whose banking system was much less regulated than that of the United States. In particular, Canadian banks had always been free to branch nationwide and were able to produce currency without bond collateral requirements. Until the early 20th century, this system was as close to laissez-faire banking as any other in the world at the time.[6] Its success was clear: Canada did not have the history of bank panics and failures that characterized the United States before World War I. The importance of Canada's lack of branching restrictions became even clearer during the Great Depression. Almost 10,000 US banks failed during this period, the vast majority of which were small "unit banks" that were highly underdiversified. Those bank failures were a chief cause of misery for much of the population during the Depression.

By contrast, no Canadian banks failed after 1929, thanks to their higher degree of geographic diversification leading to less risky portfolios. Some banks closed branches, but that only meant inconvenience for customers, not a loss of their savings. Canada escaped the worst effects of the Great Depression precisely because its banks were less regulated and because its lack of a central bank meant that it avoided both the artificial boom of the 1920s and the severe deflation of the early 1930s that caused so much suffering in the United States. The clearly destructive role of regulations in the United States and the success of the very free Canadian system show how the argument that laissez-faire in the United States caused the problems that necessitated the creation of the Fed does not hold together.

Whatever the causes, the problems were real. After 1893 and especially after 1907, debate over alternatives heated up. The turn of the 20th century marked the period known as the Progressive Era, during which the federal government acquired a variety of new powers. The prevailing view was that an administrative state guided by wise and well-meaning bureaucrats could remedy the problems supposedly created by capitalism. By 1907, the federal government had already passed antitrust laws, the Pure Food and Drug Act, and numerous other regulations. Into this context came the debate over monetary institutions. The dissatisfaction with the national banking system led to a series of commissions and debates discussing reform. After the 1893 panic, there was much debate about removing the bond collateral requirements; doing so was a cornerstone of a reform report produced by the Indianapolis Monetary Commission in 1898. However, the time was not yet right politically for significant reform.

After the Panic of 1907, a number of bills came before Congress that would have effectively addressed the problems by removing the bond collateral requirements

6. Canada did not even have a central bank until 1935, although the federal government did expand its role in currency creation somewhat after World War I.

and permitting some degree of interstate branching. Ending the bond collateral requirements was politically acceptable, but opening up branching was strongly opposed by smaller, agricultural states that feared that big city banks, especially New York ones, would enter their markets and drive them out of business. With each state having two votes in the Senate, these economically wise attempts at reform died a political death.

The states needed to reach a compromise, and all along, some bankers and groups of political activists, especially those associated with farmers, had been arguing for some sort of central bank. These arguments were consistent with the general Progressive Era belief that wise, well-intentioned, and well-trained citizens could use the power of the government to ameliorate the problems caused by unregulated markets. The reformers arguing for a central bank believed that the most serious problem was making sure that reserves could be moved to where they were needed rather than being more centralized in New York and a couple of other major cities. Bringing some sort of intentional cooperation to reserve management could be accomplished through a central bank. However, the same rural areas that feared big city banks with the ability to branch also feared a central bank that would be tied too closely to those same banks or to Washington.[7] As negotiations between politicians and bankers continued in the early 1910s, the two groups eventually settled on the Federal Reserve System as a "decentralized central bank." By carving the United States into 12 districts and putting a regional bank in charge of each one, with all 12 coordinated by a weak oversight board in Washington, DC, the hope was to diffuse the political opposition. The regional banks were also nominally private, as they would be "owned" by banks in their district that became members of the Federal Reserve System. Though technically private, each of the 12 banks had extensive government privileges—including the exclusive right to issue currency—and the ownership shares that member banks owned could not, and cannot, be traded.

In no sense is the Fed some kind of ideal central bank that economists would draw up on a blackboard to implement.[8] Like almost every other form of government activity, including the activities of the rest of the Progressive Era, the Fed was a product of political compromise in which the special interests of various parties, rather than some overarching agreement on the public interest, drove the eventual outcome. Politics inevitably produces regulations and institutions that reflect the interests of the private parties involved in the process, and the Fed is no different. That the Fed has not performed very well over its 100 years should therefore come as no surprise. The next section details how the new Federal Reserve System's structure evolved over the 20th century and how it interacted with monetary policy.

7. The United States has a long history of fearing centralized monetary power, whether private or public.
8. Historically, most central banks came into existence not because of demonstrated "market failures," but due to the government's need to raise revenue, typically for war.

HOW THE FED ATTEMPTS TO CHANGE THE MONEY SUPPLY

The conventional story of how the Fed attempts to change the money supply emphasizes the three tools it has it its disposal: the reserve ratio, the discount rate, and open market operations. The reserve ratio refers to the percentage of cash and deposits that the Fed's commercial member banks must keep on hand to redeem their customers' checks or other withdrawals. Today, that ratio is about 10 percent. The discount rate is the interest rate that the Fed (technically, each of the district banks) charges banks that wish to borrow reserves directly from the Fed. Open market operations, the Fed's most commonly used tool, occur when the Fed purchases or sells government bonds as a way to increase or decrease the quantity of reserves that banks have. When banks have more reserves, they can create more loans and expand the money supply, and when their reserves fall, they can create fewer loans and the money supply contracts. Understanding exactly how these tools work requires that we first understand how commercial banks work and that we have a little bit of historical context.

Central banks do not usually create money directly. The Fed can authorize the Treasury's Bureau of Engraving and Printing to print up Federal Reserve notes and ship them to banks, or, in theory, purchase goods and services with them, but money creation does not usually happen this way. Instead, individual banks create money in the form of the loans they make to their customers. Banks' ability to make those loans depends on the quantity of reserves they have, either as cash in their vaults or, mostly, on deposit at the Fed. Because those deposits at the Fed are in the form of bookkeeping entries, not physical stacks of bills, a central bank with a monopoly over currency production can write IOUs off of itself without limit. The Fed can simply create reserve balances out of thin air because it can, in theory, always print the currency to back up those balances should banks wish to "cash out." Put differently, the Fed can do the equivalent of writing a check off itself to give banks reserves as loans or to pay for the bonds it purchases.

An individual bank's ability to make new loans depends upon its level of excess reserves, so when the Fed adds to banks' reserves, banks can make additional loans. Banks, as noted earlier, have to hold about 10 percent against their outstanding liabilities. In addition, banks might wish to have a little bit extra on hand at times. Anything banks have in their reserves beyond that 10 percent, however, can be lent out to earn interest. Banks create those loans by creating additional account balances for the borrower, who then has the money to spend. Again, banks accomplish this process through bookkeeping entries. As long as a bank has more reserves, it can create more liabilities against them. Those additions to borrowers' deposit accounts are also additions to the money supply. Thus, it is banks that actually expand the money supply, although the Fed can make such expansions more likely by adding reserves to the banking system. The central bank does not directly expand the money supply; it does so indirectly via its manipulation of reserves.

The reserve ratio and, in particular, the discount rate were the standard tools by

which central banks influenced the money supply for centuries. One of the most important functions of central banks is to serve as a "lender of last resort" by providing liquidity to banks in times of crisis. If banks run low on reserves and customers are cashing out their deposits, central banks can provide those reserves by creating them out of thin air. However, such lending is supposed to be true lending, and the interest rate that central banks charge on such loans can be adjusted to encourage more or less of it. Lower rates would, on the margin, mean more borrowing, and higher rates less. As banks borrow more, they increase their reserves, which can be used to provide liquidity in times of crisis or can be turned into loans and increases in the money supply in normal times. The problem with the discount rate as a tool is that banks must still make an intentional decision to borrow. The Fed can only change banks' incentives; it cannot force them to borrow.

Similarly, central banks can adjust the reserve ratio to create or subtract reserves. If the reserve ratio is increased, more of each bank's reserves become legally required, meaning fewer are excess, restricting its lending. If the ratio is raised enough, banks will actually have to shrink their outstanding loans by either calling in loans or not relending as other loans are paid off, to get their reserve-to-deposit ratio where it needs to be. If the ratio is lowered, however, banks will suddenly find themselves with excess reserves to lend.

The reserve ratio is not normally used as a tool for policy because its effects are so large. It does not provide the sort of surgical precision that central banks believe they need for short run adjustments to the money supply. Banks may also mostly avoid reserve requirements through the use of sweep accounts, which involve banks taking the funds from standard checking accounts and transferring them overnight to savings accounts, which are exempt from reserve requirements, and then transferring the funds back the next morning.

Congress has only changed reserve requirements a handful of times in US history, and in all but one case, the changes were not made as a matter of monetary policy. They were made to adjust to institutional and technological changes in the banking system, or to simply increase bank profits by providing banks with zero-cost reserves.[9] In the former cases, such as the Depository Institutions Deregulation and Monetary Control Act of 1980, the expansionary effects of dropping reserve ratios were countered with contractionary policy elsewhere. In the United States, the Fed is also bound by limits to those ratios that Congress has set. Several other countries, including Canada, have eliminated reserve requirements altogether in recent decades, with no adverse effect on their ability to control the money supply.

9. The exception was the increase in reserve requirements in 1936–37, which the Fed thought was necessary to prevent inflation due to an accumulation of excess reserves. Banks responded by holding even more reserves to keep their desired level of excess reserves, leading to a reduction in the money supply. For many years, this response was thought to be the primary cause of the subsequent "recession within the depression," but more recent scholarship has suggested that regulatory factors, especially changes in labor law, might have played a larger role.

Given the problems with both changing reserve requirements and the discount rate, most short-run monetary policy is conducted through open market operations. Open market operations occur when the Fed buys or sells US government bonds as a way to change the supply of money. When it buys bonds it pays for them with newly created money, and when it sells them it destroys the money that is used to purchase them. In the United States, the Fed used open market operations very sparingly during the 1920s and then only by the New York district bank. At the onset of the Great Depression, the New York bank made somewhat more aggressive use of this tool, but not effectively enough to prevent the calamitous drop in the money supply that turned a recession into the Great Depression.[10]

Open market operations became the major tool of monetary policy starting in 1935. The Banking Act of 1935 changed the Fed in two major ways. First, it replaced the old Federal Reserve Board with a Board of Governors and then created a separate Federal Open Market Committee (FOMC) to oversee open market operations. The FOMC's members include the seven governors plus the New York Fed president and three other district bank presidents in a rotation. This group meets regularly to conduct what is normally called "monetary policy." Prior to 1935, the individual regional banks could conduct independent monetary policy. The new structure centralized the process in Washington.

The FOMC decides on the various targets for monetary policy. One misconception is the idea that the Fed "sets" interest rates. It does not, other than its own discount rate. When reports say that the Fed will keep rates low or change them in some way, they are really referring to the Fed's interest rate *target*. Normally, the major target reported is the federal funds rate, which is the interest rate banks charge each other when they lend and borrow reserves (as opposed to doing so from the Fed). The Fed's actions strongly influence the federal funds rate, which is why it is a target, but the rate itself is ultimately set by the market for excess reserves between banks. The Fed's targeted rate affects other interest rates in the economy since many of those other interest rates are linked to the Fed's targeted rates, but the Fed does not "set" rates as if it had the power of a price control.

When conducting open market operations, the Fed buys and sells various government bonds, mostly with a group of authorized security dealers. When the Fed buys bonds from them, it pays for those bonds with newly created reserves. Those reserves are credited to the bond dealer's bank's reserve account at the Fed, and that bank in turn credits the bond dealer's account with them. Again, no physical money changes hands—banks effect these transactions through a series of bookkeeping entries. The Fed's ability to create liabilities off itself out of nothing is what drives this whole process. The result of the bond purchases is that the bond dealers have swapped one asset (government bonds) for another asset (cash) and the banks have

10. Milton Friedman and Anna Schwartz, *The Great Contraction: 1929–1933* (Princeton: Princeton University Press, 2008).

gained both reserves and new deposits. The Fed has created additional reserves and acquired government bonds, almost all the interest from which is returned to the Treasury. By returning that interest, the Fed is essentially giving the federal government an interest-free loan every time it buys government bonds. Thus, it is possible for the government to create new debt without incurring more interest costs.

Not only do open market operations affect the money supply, they also matter for fiscal policy. The power of central banks to purchase government debt by creating money out of nothing allows governments to run larger annual deficits and accumulate more total debt than they would be able to otherwise. If the public does not wish to buy government debt, the central bank can always pick up the unpurchased supply. The ability to use a central bank to finance government expenditures has long been a rationale for creating such banks in the first place, or for expanding their powers in ways to make such finance possible. Historically, this process was most often associated with the revenues for fighting wars, especially unpopular ones, but as we have seen with the Fed's actions since the 2008 financial crisis, those powers can fund all kinds of deficit spending. The danger of this interaction of fiscal and monetary policy is that an insufficiently independent central bank can easily begin to monetize larger and larger amounts of government debt, potentially leading to high rates of inflation. We have seen this happen in countries like Brazil, Israel, and Zimbabwe in the last few decades, and there is no reason why it could not happen to larger Western economies.

The US economy has seen notably higher average rates of inflation over the later years of the Fed's existence, as well as much higher government debt. One other change affecting monetary policy has influenced this outcome. In the Fed's early years, Federal Reserve notes were redeemable for gold, and gold coins still circulated in the US economy. That practice ended in 1933 when President Roosevelt took the United States off the domestic gold standard, though foreign central banks could redeem Federal Reserve notes for gold. During the 1960s, inflation rates began to rise under Presidents Johnson and Nixon, partially as a result of using newly created money to finance the Vietnam War. As those newly created dollars flowed overseas, foreign central banks began to redeem them for gold, leading to concerns about the US gold stock. In 1971, President Nixon decided to end redemptions in gold for foreign central banks ("close the gold window") in order to preserve the gold stock. The result was the end of the last, if weak, economic check on the Fed's ability to inflate. Until Nixon's action, the Fed at least had to consider the cost of losing gold when foreign central banks redeemed Federal Reserve notes. After 1971, that concern was gone, and the inflation rates of the late 1960s became even worse in the 1970s and '80s, leading to some of the worst economic times since the Great Depression.

In the wake of the 2008 financial crisis, the Fed began to exercise a variety of powers that it had not made use of previously, most of which involved providing direct credit to financial institutions. These new powers were part of the Fed's

discretionary assertion of a responsibility to prevent financial institutions from failing. The Fed has always had some regulatory powers, but during the crisis, it seemed to see itself as the only institution capable of acting in the way it perceived was necessary. As a result, the Fed saw no need to ask for congressional approval to expand its powers beyond those it had historically exercised.

Of these new powers, two are especially relevant to the conduct of monetary policy. First, almost immediately after the onset of the crisis, the Fed decided to begin to pay interest on the reserve balances that banks have with it. The amounts have been small, but this change was nonetheless a first, and it has given the Fed what amounts to a fourth tool, known as the "corridor system." The corridor refers to targeting the federal funds rate in such a way as to be between the floor, or the interest rate being paid on reserves, and the ceiling, or the discount rate. Now that the Fed can set the interest rate on reserves and the discount rate, this combination has become the new major tool for monetary policy as the Fed moves forward.

The most obvious explanation for the Fed's decision to start paying interest on reserves is that as long as that rate is above the risk-adjusted yield on other assets, it will have a contractionary effect on the money supply by encouraging banks to hold rather than lend their reserves. Discouraging bank lending in this way was almost certainly the rationale for this policy change, because it occurred simultaneously with the Fed injecting capital into failing banks. Those capital injections were not only in violation of Bagehot's principle of "lend freely to healthy banks at penalty rates," but they also ran the risk of igniting inflation. Paying interest on reserves helped offset that threat by encouraging banks to sit on the new reserves. The conjunction of these two policies also suggests that the clear purpose of those capital injections was to ensure bank solvency, not bank liquidity. That is, the point was to rescue failing banks, not to expand the money supply. If the goal were the latter, why offset the capital injections with interest payments on the reserves they created?

The second new aspect to monetary policy came a little bit later in the fall of 2008 in the form of "quantitative easing." Quantitative easing is a form of expansionary open market operation in which the Fed buys financial assets on a significantly larger scale than it does during standard open market operations. As part of its quantitative easing strategy, the Fed has purchased long-term financial assets rather than the short-term government bonds typically associated with standard open market operations. Quantitative easing is also distinct from standard open market operations in its focus on adding a specific quantity of reserves into the banking system rather than seeing changes in reserves as a means to a further interest-rate or money-supply targets. Starting in November of 2008, the Fed has engaged in three programs of quantitative easing. It has bought a variety of financial assets, including longer-term government bonds and mortgage-backed securities, from banks and other financial institutions.

In September of 2012, the Fed announced a shift, generally referred to as "QE3," in its quantitative easing strategy. What distinguishes QE3 from its predecessors is

that it is an open-ended commitment to buy $40 billion worth of mortgage-backed securities and $45 billion worth of longer-term Treasury securities per month for the foreseeable future. Earlier rounds of quantitative easing were not open-ended and were less focused on longer-term assets. As part of QE3, the Fed also announced that it would continue to keep interest rate targets extremely low well into 2015. Toward the end of 2012, the Fed went one step further and announced that it will keep the federal funds rate near zero as long as the unemployment rate remains above 6.5 percent, and its projections of inflation one to two years out are no higher than 2.5 percent. The Fed's use of dated policy, such as committing to keep interest rate targets low into 2015, is referred to as "forward guidance," reflecting the way in which it informs the public of its policy intentions. The addition of the unemployment and inflation criteria influencing the Fed's forward guidance helps to dispel the idea that the Fed intended to keep interest rate targets low regardless of economic conditions. The combination of the open-ended purchases and the extension and expansion of forward guidance represents a new policy stance for the Fed.

Quantitative easing has been adopted in countries such as Japan where short-term interest rates have fallen near zero and where central banks saw clearing junky assets off bank balance sheets as a worthy goal. The hope is that by increasing the quantity of reserves, the policy will lead to lending and spark growth, in part by bringing down longer term interest rates. However, this outcome depends on banks having viable lending options in a low interest rate and low-growth economy, and it requires that those options are superior to the rate the Fed is paying on those reserves. Four years into quantitative easing, most of the reserves created are still sitting in banks' accounts at the Fed, costing the Fed more in interest with each round. Unorthodox monetary policy will not work when the rest of the economic environment is not conducive to expectations of good returns on loans.

QE3 raises a number of other concerns. As with all expansions of bank reserves, the most serious danger is inflation. The combination of interest payments on reserves and a dearth of safe investment opportunities in the market has led banks to accumulate massive quantities of excess reserves, preventing much of the expansion of the monetary base from entering the spending stream in the form of an excess supply of money. The Fed keeps pouring in liquidity, but the size of the vessel just expands, preventing reserves from spilling over into the spending stream. This expansion cannot go on forever. Should the yields on alternative assets rise, perhaps due to widening real recovery, unless the Fed is willing to raise the interest rate on reserves and the Treasury is willing to bear the cost, that interest rate will be insufficient to keep reserves in the banks. To some degree, the Fed wants those excess reserves to enter the spending stream to facilitate the consumption and investment spending believed necessary to increase GDP and reduce unemployment, but the challenge is to find a viable exit strategy that will get the excess reserves out of the banks before too many are lent out, leading to inflation. If the Fed cannot do so, inflation may be a significant problem down the road.

In the short run, QE3 continues to be a way to bail out financial institutions with underperforming mortgage-backed securities. Fed purchases of such assets have enabled banks and others to get these assets off their books in exchange for reserves offering a very small but still positive rate of return. To that degree, QE3 is a de facto bailout program, and an open-ended one at that. In addition, the Fed's determination to keep interest rates near zero reduces the incentive for households to save, which is the opposite of what the US economy needs after the destruction of so much capital in the housing boom and bust. Sustainable economic recovery requires real savings by households in order to provide the lending necessary to rebuild the capital lost during the recession, and QE3 undermines the incentives for saving by continuing to target interest rates near zero.

ECONOMIC THEORY AND MONETARY POLICY'S OBJECTIVES

Understanding the tools at the Fed's disposal is only the first step in understanding monetary policy. What the tools do not tell us is what we should use them to do. That is, what are the goals of monetary policy, and what instruments should guide us to those goals? Should the central bank attempt to control inflation? Should it try to influence unemployment? Should it attempt to boost measures such as gross domestic product? Given our answers to those questions, which of the variables that the Fed can influence should it be targeting in order to achieve those larger policy goals? Are interest rates the key? The money supply? The issues here are at the heart of macroeconomics and are among the most contested in the discipline of economics. The literature on these issues is much deeper and more complex than this discussion can do justice to, but in this section, I present an overview of some of the different macroeconomic positions on these topics.

In their early years, central banks did not really engage in monetary policy as we now understand it. Central banks were largely charged with raising revenue for governments and with being lenders of last resort in times of crisis. These two goals have an interesting relationship, as generating inflation to create revenue often caused the very crises that central banks were supposed to either prevent or resolve. As the theory of money progressed into the early 20th century and management of monetary policy became part of that work, the general guideline for central banks was to focus on price-level stability. The assumption was that some aggregate measure of the price level could serve as a sufficient guide to avoiding the problems of inflation and deflation. Much of the early 20th century literature discussed the reasons why price-level stability would promote a sound environment for economic growth. The work of American economist Irving Fisher is a good example of this way of thinking.[11]

In Europe, several thinkers were developing a somewhat different approach

11. Irving Fisher, *The Purchasing Power of Money* (Fairfield, NJ: Augustus M. Kelley, [1913] 1985).

to the ideal monetary policy, foremost among them economist Knut Wicksell of Sweden.[12] Wicksell argued that interest rates were the key indicator of sound monetary policy. He distinguished between what he termed the "natural" rate of interest and the market rate of interest. The latter was the rate actually being charged for loans in the market, while the former was an analytical construct representing the underlying time preferences of lenders and borrowers. If the market rate equaled the natural rate (i.e., if the market signal accurately reflected people's real preferences), prices would be stable and monetary policy would be doing its job. Any deviation between the two rates would result in inflation or deflation (if the market rate were lower or higher than the natural rate, respectively). The challenge of using Wicksell's approach as a policy guideline was that the natural rate is unobservable, so there was no way to look at a specific market interest rate and know for certain whether it was matching the natural rate. Instead, the question was whether some sets of monetary institutions would be better than others at keeping those two rates close by ensuring that the public's saving, and not central bank injections of new money, were the only source of funds for investment.

Wicksell's work was the basis for a theory of the business cycle developed by the Austrian school economists Ludwig von Mises and F. A. Hayek in the decades before the Great Depression.[13] They took Wicksell's insight about the natural and market rates and combined it with a theory of the capital structure of the economy to help explain the historical observation that artificial booms tended to be characterized by growth in capital goods far from the final product consumers purchased, such as the machines that produce parts that eventually go into engines that later end up in cars. What the Austrians argued was that when the market rate fell below the natural rate, most often due to central bank inflation, it falsely encouraged entrepreneurs to invest in the early stages of longer-term production processes (e.g., research and development) because the lower rate made it look like there was more saving to finance the additional investment. This growth characterized the boom. Eventually, the lack of real savings would make itself known and would trigger the bust.

For the Austrians, unlike Wicksell, the price level was not a good guide to the appropriate policy. Their theory of capital, especially in work by Hayek in the late 1920s, helped them understand that the growth in productivity that pushed down prices was a good thing and happened in ways that did not disrupt macroeconomic stability. Rather than counter productivity-induced declines in the price level with expansionary monetary policy in order to maintain price-level stability, the Austrians recommended letting those price declines happen. For example, we

12. Knut Wicksell, *Interest and Prices* (New York: Augustus M. Kelley, [1936] 1965).
13. Ludwig von Mises, *The Theory of Money and Credit* (Indianapolis: Liberty Fund, [1912] 1980); F. A. Hayek, *Prices and Production*, 2nd rev. ed. (New York: Augustus M. Kelley, [1935] 1967; F. A. Hayek, *Profits, Interest, and Investment* (Clifton, NJ: Augustus M. Kelley, [1939] 1975). A modern version of the theory can be found in Roger Garrison, *Time and Money: The Macroeconomics of Capital Structure* (New York: Routledge, 2001).

would not want to offset the falling prices of technology in the last 20 years with expansionary monetary policy. Those declines have enabled us all to live better for less money. Hayek argued for a concept of "neutral money" in which banking institutions allowed prices to be determined on the basis of real factors of supply and demand. To achieve this goal, the banking system had to maintain what was known as "monetary equilibrium." That is, it had to supply the quantity of money that people were willing to hold at the current price level. As later Austrians have clarified, doing so would ensure that the amount of funds available for investment was equal to the public's intended saving, which in turn meant that the natural and market rates were equal.[14] The challenge here as well is how to put this theory into practice. This challenge is one reason why Austrian school economists tend to support alternatives to central banking (see below), as they believe that only a banking system without a central bank and where money is produced competitively can approach monetary equilibrium.

These approaches to monetary policy were swamped in the 1930s and 1940s by the Keynesian revolution.[15] John Maynard Keynes and the later Keynesians changed how people thought about the role of central banks in two ways. First, they emphasized fiscal policy over monetary policy. Early Keynesian economists (including Keynes himself) and most later Keynesian economists argued that it was up to fiscal policy to balance the economy by unbalancing the budget. When the economy weakened, governments should deficit-spend to boost aggregate demand to levels compatible with full employment. When the economy recovered and was going strong, governments should run surpluses to pay off the debt accumulated during the downturn. The budget would thereby be balanced across the business cycle. This line of thought put monetary policy in a secondary role, at most given the task of making sure that interest rates were low so that governments could borrow cheaply.

Later Keynesians, through what was known as the "neoclassical synthesis," gave monetary policy a more central role. It was still the relationship between money and interest rates that mattered, however. Monetary policy could play something of a role in stimulating aggregate demand because expansionary policy could drive down interest rates, which would in turn spur increased investment. That investment would work through the spending multiplier process into increasing GDP as well as employment and other relevant variables. The debate during the 1950s and 1960s was largely over two questions: (1) how much influence would a given increase in the money supply have on lowering interest rates, and (2) how much new investment would a given decline in interest rates

14. Steven Horwitz, *Microfoundations and Macroeconomics: An Austrian Perspective* (New York: Routledge, 2000).
15. John Maynard Keynes, *The General Theory of Employment, Interest, and Money* (New York: Harcourt, Brace and Company, 1936).

produce? Keynesians believed the answer to both was "not much," though it was still greater than zero. So to the extent monetary policy was effective, its focus should be on interest-rate targets. Contemporary Keynesians use different arguments and are not as skeptical of the money-interest-investment channel's ability to be effective. As a result, they think monetary policy matters more, but they continue to think that interest-rate targets are the most important tool for generating economic growth and low unemployment.

The major contending perspective on monetary policy is one that puts more emphasis on measures of the money supply and the behavior of the price level as the proper guides to good policy. The modern version of this view, eventually known as "monetarism," began with the work of Milton Friedman and others in the 1960s. Friedman argued that changes in the money supply were the cause of various macroeconomic disturbances in US history, especially the Great Depression.[16] Some Keynesians at the time argued that workers would not understand how inflation was reducing the real value of their wages, and would therefore continue to work more as inflation rose and firms had more to spend on expanding employment. Friedman believed that inflation was not capable of fooling workers into working more, at least in the long run, because workers understood the effects of inflation and adjusted their wage demands accordingly; thus attempts to use monetary policy to generate declines in unemployment and increases in GDP were likely to be useless in the long term. It is the real variables in the economy, such as tastes, technology, resources, and good institutions, that promote sustainable long-run growth. Friedman's historical evidence and his revision of theory forced economists to take more seriously the role of the money supply and the importance of the price level in monetary policy.

Changes in practice in the late 1970s and early 1980s, when the Fed briefly shifted from its longstanding sole emphasis on interest rates to take more substantial account of monetary aggregates, reflected Friedman's influence. Though people still debate whether the Fed was really following Friedman's advice, after a few years it shifted policies again. By the mid-1980s, both interest rates and money supply targets were part of the policy making process, as they largely remain today.

The final part of this debate is the role of the Fed's "dual mandate." Consistent with early 20th century thinking, the Fed's original mandate was to maintain stable prices. However, with the Keynesian revolution and the change in thinking about monetary policy, the Fed's mandate was amended in 1977 to include a reference to promoting "maximum employment." For the most part, this mandate has been interpreted to mean the maximum employment compatible with stable prices, but since 2008, more people associated with the Fed have been using the language of

16. Milton Friedman and Anna Schwartz, *A Monetary History of the United States, 1867–1960* (Princeton: Princeton University Press, 1963), and Milton Friedman, "The Role of Monetary Policy," *American Economic Review* 58, no. 1 (March 1968): 1–17.

"full employment." The challenge for the Fed remains trying to keep prices stable and promote full employment, as many politicians and economists believe that during downturns, expansionary monetary policy is necessary to generate full employment. This idea has certainly guided the Fed during the recession and slow recovery of the last few years. Those of a more Keynesian bent have welcomed the renewed focus on full employment, while followers of Friedman and the modern "monetarist" tradition (as well as Austrians) have raised concerns about the inflationary threat posed by sacrificing price stability for what they see as the chimera of full employment. For the latter groups, monetary policy, like medical professionals, should "first, do no harm." With the Fed having acquired new powers and taken on new tools, the critics of the dual mandate who are concerned about price stability have been more vocal in recent years.

MONETARY DISEQUILIBRIUM: THE DANGER OF BAD MONETARY POLICY

If monetary equilibrium (i.e., producing the quantity of money that the public wishes to hold at the current price level) is the goal of monetary policy, then the danger is that bad monetary policy will create monetary *disequilibrium*, producing more or less money than the quantity the public wishes to hold. An excess supply of money is called *monetary inflation*, and a deficient supply of money is called *monetary deflation*. These definitions are important because an excess or deficient supply of money may or may not be associated with increases or decreases in the price level.

One of the arguments made by those who favor monetary equilibrium as the appropriate norm for policy is that the overall level of prices can be affected not just by too much or too little money, but by changes in productivity as well. In a growing economy with rising productivity, the long-run tendency will be for prices to fall over time as goods become cheaper to produce. So, for example, an excess supply of money may not lead to rising prices if the upward pressure on prices is more than offset by the downward pressure of rising productivity in a growing economy. Movements in the price level due to changes in productivity are fine, but changes arising from monetary disequilibria are not. We want prices to change when the supply and demand for specific goods and services change, or when economy-wide productivity changes, not when there is too much or too little money. For the remainder of this paper, I will use the words "inflation" and "deflation" to refer to monetary disequilibria involving an excess and deficient supply of money, respectively. Where I refer to changes in the price level, I will modify those terms with the adjective "price."

As an empirical matter, central banks are far more likely to err on the side of creating too much money than too little money because, as discussed earlier, governments benefit when central banks expand the money supply. Open market operations clear room for governments to issue more debt for roughly the same interest

cost. Unexpected inflation, as we will see, benefits borrowers and hurts lenders by reducing the real value of debt. With governments being tremendous debtors, inflation can provide an irresistible path toward debt "relief." As a result, the independence of central banks is crucial. The more they are under the direct sway of the legislative or executive branches, and especially the Treasury, the more likely they will be to use excess supplies of money to help reduce the costs of borrowing. By contrast, deflation has all the opposite effects for highly indebted governments. It *increases* the real value of debt and punishes governments that borrow. Central banks, and especially ones that are closely tied to governments, will normally not intentionally engage in deflationary monetary disequilibria. When deflation does happen, as it did in the United States from 1930 to 1933, it is usually the result of error, not intentional policy.

In understanding the effects of monetary disequilibria, we need to note that people wish to hold a particular quantity of real purchasing power in the form of money, which is our demand for money. We demand money when we hold it. However, we can find ourselves with less or more money than we wish to hold. That is, our *actual* holdings of money can be less than or greater than our *desired* money holdings. A deficient supply of money causes our actual holdings to be less than our desired holdings, and an excess supply causes the reverse.[17] The effects of each are the result of the actions we take to bring our actual and desired holdings back to equilibrium.

Monetary Disequilibrium I: Too Little Money

The effects of an actual money supply that is less than what people desire to hold are usually obvious and quickly felt. Modern economies need a sufficient supply of money to ensure that as many mutually beneficial exchanges as possible get made, much like an engine needs oil to ensure it runs most efficiently. Without a sufficient money supply, economies, like engines without oil, grind to a halt.

Consider the options facing the person whose actual holdings of money are not as much as she desires. To get more money, she can do any of the following:

1. earn additional income from more hours or an additional job
2. sell off some other assets for cash
3. restrict her expenditures so that more of her income stays in her cash balances

All three can work, but the first two are less likely to succeed because they depend on other people choosing to pay her, which implies that those people *have excess money they wish to part with*. If the central bank has produced an insufficient money

17. It is important to not confuse "wealth" and "money." Perhaps we can never have too much wealth, but we can have too much of our wealth *in the form of money*. When we purchase goods and services, we are indicating that we'd rather have those forms of wealth than the money form.

supply, a large percentage of the population will have insufficient cash balances, which rules out the first two options.

The effects of the third will be to slowly bring economic activity to a halt. If I restrict my spending to replenish my cash balances, that means lower income for those on whom I would have spent my money. They will in turn see lower cash balances than they want, leading to them reducing their spending, and so on. The cumulative result will be a slowing down of economic activity. Producers will see sales fall and will start to reduce inventories, shed employees, and purchase fewer inputs. The increase in unemployment will magnify the problems as those people lose their incomes. The contraction in demand and income is not due to any reduction in the real resources available in that economy, as its human and physical capital is all still there. Rather, without money, the economy cannot turn those resources into mutually beneficial exchanges.

The problem will solve itself over time, but not without economic pain. Assuming the central bank takes no action to raise the money supply back to where it should be, producers will eventually realize they have to cut their prices to attract consumers who hold lower money balances. As prices fall, households will be less unhappy with the gap between their actual and desired money holdings, since they do not need the same amount of nominal money balances with lower prices. Spending will pick up and employment and input purchases will climb as well. However, this adjustment process can take many months or even years, during which unemployment will be high and investment will be low.

This is more or less what happened during the early 1930s, though the process was complicated by President Hoover's unfortunately successful attempt to convince major industrial firms to *not* reduce nominal wages in the face of what ended up as a 30 percent decline in the money supply. By maintaining nominal wages at the predeflation level, Hoover and the industrialists prevented the downward wage adjustments that were needed to offset the fall in prices caused by the drop in the money supply. The result was that unemployment quickly rose to almost 25 percent by 1932. The pain caused by a deficient money supply is problem enough, and Hoover multiplied the pain by jawboning stickier wages.[18]

Unexpected deflation also favors lenders over borrowers. As the price level eventually declines, and the value of each dollar thereby rises, the borrowers end up paying their loans back in dollars that are more valuable than those they borrowed. In other words, the real value of their debt rises.

Finally, we should distinguish the fall in prices due to an insufficient money supply from the fall in prices due to gains in productivity. When producers are able to find new and better ways of making their products, leading to declining prices over time (as we often see with new technology), that is an unambiguous improvement for households and the entire economy. Such productivity-induced declines in the

18. Steven Horwitz, "Herbert Hoover: Father of the New Deal," Briefing Paper No. 122 (Washington, DC: Cato Institute, September 29, 2011), http://www.cato.org/pubs/bp/bp122.pdf.

price level are not a cause for concern, nor should monetary policy try to offset them in order to "stabilize" prices. Doing so would require creating excess supplies of money and would generate all the problems discussed in the next section. Prices should be allowed to change due to real factors in the economy and should not be offset by policy. Productivity-induced declines in prices are part and parcel of healthy economic growth and should not be confused with downward pressure on prices caused by deflationary monetary disequilibria.

Monetary Disequilibrium II: Too Much Money

The costs of inflationary monetary disequilibria are much more subtle and complex than those of deflationary monetary disequilibria.[19] Unlike deflation, which shows up very quickly as *idled* resources, making it very clear that there is a problem, inflation instead leads to *misallocated* resources, the damage from which is more hidden and long-term. Where an insufficient money supply means that mutually beneficial exchanges simply never get made, an excess supply of money means that the exchanges that *do* get made are not the ones that would be made without the inflation. The pattern of exchanges caused by an excess money supply creates misallocations and distortions that ultimately undermine economic growth.

Understanding these costs of inflation requires recognizing that excessive money creation does not cause all prices to rise by the same amount. In a mirror image of a deficient money supply, an excessive supply of money causes people to have actual cash balances that are *greater* than they desire, which will lead them to get rid of the excess by spending it. Those expenditures are what put upward pressure on prices. The problem is the excess money supply; the upward pressure on prices is just the visible manifestation of the inflation. When people spend, their own preferences and trade-offs will guide their choices, and some prices might go up by a lot and others by only a little. Some could even fall. The *average* level of prices will be higher than it would otherwise thanks to the increased spending, but individual prices will change in a variety of ways. These movements in individual prices matter because what consumers and producers care about are the relative prices of goods, or the price of one good compared to another. If all prices rose by the same percentage, relative prices would be the same and the allocation of resources would not change. When prices increase by varying degrees, those changes pose problems for producers and consumers.

People already have to deal with the daily challenge of interpreting price changes, and the movements in prices caused by too much money add another layer of complexity. Now producers must determine not just what underlying supply and demand factors might have caused the prices of their inputs or products to change, but also whether and to what degree monetary disequilibrium might be the cause.

19. Steven Horwitz, "The Costs of Inflation Revisited," *Review of Austrian Economics* 16, no. 1 (March 2003): 77–95.

If the cause is a real shift in consumer preferences, for example, the firm might wish to add workers or acquire a new machine. If the cause is the temporary influence of the excess money supply, the firm might not wish to expend resources on reacting, as those changes could be reversed in the near future. The problem is that price changes do not come stamped with "monetary policy" or "consumer preferences" or any other cause of their movement. Producers must *interpret* them. What inflationary monetary disequilibria do is make that process unnecessarily more complex, making it more likely that producers will make mistakes and misallocate resources in the process.

One of the most fundamental insights of economics is the role that prices play in resource allocation. Prices serve as knowledge surrogates that help producers and consumers coordinate their choices with those of others in the market. Prices are, in this way, bits of knowledge wrapped in incentives. When a price changes, it signals to market participants that some other variable has changed and simultaneously provides them with an incentive to change their behavior accordingly. When the price of gas rises, it *informs* us that the resource is more scarce and gives us an *incentive* to reduce our consumption to conserve it. Monetary disequilibria interfere with this process by introducing the equivalent of static into the communication process, thereby making it harder for people to interpret correctly what the price change is signaling. The result is that more mistakes get made, and resources get misallocated in a variety of ways.

This misallocation is especially problematic when the mistakes involve creating new capital goods, as they cannot be used for other purposes without cost when it becomes clear that the demand for them was not really there. One example of this phenomenon is what happened during the housing boom and bust. Americans overbuilt houses in response to interest rates and housing prices being distorted by the Fed's overly expansionary monetary policy. When that boom eventually ended, owners abandoned many of those houses, and construction companies left others half-finished. Those houses represent misallocated resources (or "malinvestment"), and they cannot be costlessly converted into the resources that *are* in demand. It is costly to tear down houses and make something else out of the land or salvageable parts.[20] Thus, inflation does not just lead to mistakes that need to be corrected; it can lead to mistakes that are expensive to correct.

When it is unexpected, an excess money supply also redistributes wealth from lenders to borrowers. Without having built an expectation of inflation into the interest rate on a loan, the borrower will pay back the loan in dollars that are increasingly worth less as the excess money supply continues to put upward pressure on prices.

20. The same point can be made about investments in human resources. We had too many people in real estate and finance during the boom, and they cannot retrain themselves for different jobs without cost. The misallocation of human resources is one reason why unemployment rises during the bust. If retraining is sufficiently costly, or if there is much uncertainty about what unemployed workers should retrain *for*, high unemployment can linger.

Effectively, this inflation reduces the burden of the borrower's debt as his or her income rises while continuing to make the same nominal payment to the lender. Inflation is one reason that mortgage lenders developed adjustable rate mortgages in the late 1970s.

Through the same process, inflation will hurt small savers who have their funds in accounts not paying market rates of interest. For those who still keep their savings in cash in their homes, inflation will destroy its value. Because measures of inflation are always based on the prior year's increase in prices, those living on fixed incomes such as Social Security will not see adjustments in their benefits until after prices have actually risen, which reduces the real value of each year's payments. In all these ways, inflation redistributes wealth, most often in ways that harm the lower and middle classes.

Inflationary monetary disequilibria, or even the expectation of them, waste resources by inducing people to cope with their consequences by expending resources. For example, both producers and consumers might consult a financial adviser to help manage their assets better. Parties to contracts will have to spend more time figuring out the effects of inflation on prices and writing their contracts accordingly. Others will spend much more time watching what the Fed is doing and trying to determine how its actions might be affecting their own particular markets. All these activities might be "rational" in the sense of being appropriate responses *given the existence of the excess money supply*. However, if monetary policy were maintaining monetary equilibrium, those resources could be used to purchase or create goods and services that people directly want. It makes sense for people who live in high-crime areas to spend money on locks and security systems, but if the crime rate were lower, they could use all that money to purchase food, clothing, or something else. In this comparative sense, the threats of both monetary disequilibrium and crime waste resources that could add to human happiness if those threats were absent.

Finally, an excess money supply can trigger the boom and bust of the business cycle. I previously discussed the Austrian business cycle theorists of the interwar years. They, along with their modern descendants, argue that inflationary monetary equilibria can drive market rates of interest down below the natural level compatible with sustainable growth. In doing so, they encourage people to invest in longer-term processes of production. The lower interest rate makes it look as if there is more saving taking place and as if consumers are more willing to wait to purchase the goods they want. When more saving is taking place, interest rates fall and producers respond to that interest rate signal by taking more time to create things. Those longer-term processes are more productive, meaning that with more time, more goods can be produced. Hence, an increase in real saving leads to economic growth.

However, when excess money creation makes it look like there is more saving, producers respond as if there were more saving by adding capital and hiring more

labor. The result is a boom. When the saving is an illusion, this boom is not sustainable because the expectations of producers and consumers are not coordinated. Producers expect delayed consumption from the lower interest rate, but consumers *have not in fact saved more* and do not wish to wait. So while producers have planned to have output available farther in the future, consumers want things sooner. This disjunction cannot be sustained. Eventually, the mistakes of the boom become apparent, and we get a bust. The cause of the business cycle is the monetary disequilibrium having driven interest rates artificially low and created the incentive to invest in the unsustainable long-term projects that constitute the boom. The waste represented by the mistakes induced by the boom is a significant problem created by poor monetary policy.

The costs of both kinds of monetary disequilibrium are significant, which is why making sure that the institutions responsible for supplying money maintain monetary equilibrium is so important. The record of central banks in this regard is not good. In the years since the creation of the Fed, the US economy has suffered more bouts of significant monetary disequilibria, as well as more variance in GDP growth, than in the years before.[21] The housing boom and bust of the beginning of this century has further called into question the ability of the Fed and other central banks to manage the money supply effectively. Central banks across the world, but especially the Fed, have been under increasing critical scrutiny, and never has there been more serious discussion of alternatives to central banking than during the last few years.

ALTERNATIVES TO CENTRAL BANK DISCRETION

The failures of central banks to manage the money supply in ways that avoid macroeconomic disturbances should not surprise anyone. We have decades of evidence, as well as a mountain of economic theory, that indicates the futility of centrally directing resources rather than relying on markets.[22] The worst instances of inflation, deflation, and bank failures in US history have all taken place since the creation of the Fed. Central banks face a slightly less dramatic version of the same problems that face governments that try to centrally plan their economies. The problem both face is how to acquire the knowledge needed to determine what consumers want, how much of it to make, and which inputs to use to make it. For the central bank specifically, the problem is determining both the demand for money and what the relevant macroeconomic indicators are that should be guiding money supply decisions. There are two aspects to the problem. First, it takes time to collect data, analyze it, decide what to do, execute the decision, and see its effects. These "lags" are a well-known problem faced in both monetary and fiscal policy. If too much time passes between data collection and effects, the facts on the ground may have changed in ways that make the policy now taking effect inappropriate for the new conditions.

21. George Selgin, William D. Lastrapes, and Lawrence H. White, "Has the Fed Been a Failure?" *Journal of Macroeconomics* 34 (2012): 569–96.
22. Don Lavoie, *National Economic Planning: What is Left?* (Cambridge, MA: Ballinger, 1985).

Determining the demand for money is a separate problem. The central bank must learn what it is that individuals want with respect to their money holdings, but that knowledge may not be in a form that is easily accessible. As central planners across the globe have discovered, knowing the usual economic data is not the same as knowing which preferences and trade-offs are driving individual decision-making. That sort of knowledge cannot easily be put into words or statistics, making it impossible for the central planner or central bank to collect it. Without reliable information, the central bank is just guessing at that demand. As a result, central banks normally rely on the very macroeconomic data they are trying to influence as a signal that they need to act. In other words, rather than being able to get data that enables them to *avoid* a drop in GDP or rise in unemployment, it is the very observation of those changes that tells them they should act. Central banks are always playing catch-up. These problems of lags and knowledge can explain the poor track records of central banks.

Though it is less of a problem in the United States than in many countries, central banks that are not sufficiently independent can also be subject to serving the self-interest of the executive or legislative branches. Where the central bank is closely associated with the Treasury, for example, the bank can be pressured to conduct monetary policy that meets the fiscal needs of the government rather than pursuing goals more congenial to the nation's economic growth. In the United States, the Fed is officially independent of the Treasury, but until the Treasury-Federal Reserve Accord of 1951, the Fed was still informally thought to be obliged to serve the Treasury's fiscal needs. Since then, the Fed has been able to conduct policy independently. In other countries, the Treasury or even the legislature has more direct control over the central bank and can easily order it to buy up whatever debt the government creates. Such an arrangement not only encourages fiscal profligacy and enormous debt, but turns that debt into inflation by buying it up with newly created money. Central banks can fail because they cannot acquire the knowledge they need to meet their goals and because they face incentives to meet the demands of political self-interest rather than sound monetary policy.

One argument that proponents of central banks often make is that only central banks have the tools to step into macroeconomic and financial crises and solve them, and that without central banks, we would be prone to severe and lengthy crises and depressions. The problem with this argument is that it ignores the possibility that such crises are almost always *caused* by central banks. So when central banks argue that only they can solve such problems, they may be like an arsonist disguised as a firefighter who claims only he can put out the fires he started.[23] When we give central banks powers to combat crises, we should be aware that we are giving them the power to create those crises. The same open market operations that central banks

23. See George Selgin, "Federal Reserve Should Resist Tinkering," *Christian Science Monitor*, August 31, 2007, http://www.csmonitor.com/2007/0831/p09s02-coop.html.

use to supply funds in a crisis give them the ability to create the inflationary boom that leads to the crisis in the first place.

The failures of central banks to produce a macroeconomic environment conducive to growth and the repeated cycles of inflationary booms and busts have led many to ask what alternatives there might be to central banking as currently practiced. There are four alternatives to the way the Fed currently functions that might curtail its ability to create problems:

1. constraining the central bank by making the money it creates redeemable in a commodity such as gold,
2. ending the Fed's dual mandate and imposing a price-level or inflation-rate target,
3. closing the central bank and allowing banks to competitively create money and evolve their own processes for coordinating their reserve flows, or
4. targeting nominal gross domestic product.

Each of these alternatives has its own costs and benefits, and each should be the subject of serious consideration as a replacement for the discretionary monetary policy regime of central banks both here and elsewhere. I now briefly explore each in turn.

One option to fix the failures of central banks is to return to some form of a gold standard.[24] Making Federal Reserve liabilities such as currency and bank reserves redeemable in gold would place an economic constraint on the Fed's ability to mismanage the money supply, at least with respect to inflation. If the Fed were to produce more money than the public wished to hold, there would be at least the possibility of the excess money being redeemed for gold. Because banks and the Fed would want to hold a particular percentage of gold against their outstanding liabilities, redemption in gold would reduce that percentage below the desired level, forcing the banks to restrict their lending and therefore the money supply. In theory, at least, redemption demands for gold of excess supplies of money would put an economic check on the ability of central banks and the banking system to expand the money supply.

Making bank liabilities redeemable in some commodity can be effective in this way, but it is less so when done in the context of a central bank. Because even a central bank on a gold standard will maintain its monopoly over currency production, that currency will be the only option for people who wish to make hand-to-hand transactions. Gold, even in coined form, is usually more cumbersome. People's desire to use paper money makes it less likely that excess supplies will be redeemed for gold. This fact is particularly true for banks themselves, which will want to keep stocks of the monopolized currency on hand to provide to customers who prefer cash to checking account balances. The banks, too, will not have a strong incentive

24. To be precise, any commodity standard might have similar desirable effects, but for the sake of discussion, I will assume the commodity in question is gold.

to get rid of all the excess currency they might see coming their way. The institutions most likely to redeem excesses with regularity are foreign central banks that have less of a reason to hold stocks of the monopoly currency. Indeed, in the late 1960s when inflation began to appear in the US economy, foreign central bank redemptions were significant enough for President Nixon to end their right to redeem dollars for gold. By its nature as a promise by the central bank, the gold standard is also only as good as the central bank's willingness to keep the promise of redemption. Historically, central banks have reneged on those promises during times of crisis such as war and depression, especially when governments were unable to raise funds other than through inflation. Between the weaker incentives for redemption associated with central bank monopolized currency and the historical track record of broken promises, the long-run effectiveness of a gold standard is a matter of some debate. Still, even if such a change were to last for only a decade or two, it might have some significant positive impact.

The second option is to end the Fed's dual mandate by imposing a price-level or inflation-rate target. For years, many advocates of alternatives to discretionary monetary policy argued for limiting the growth of the money supply, either by internal Fed policy, by statute, or constitutionally. Changes in theory and the inability of those policies to work in practice in the face of much greater variation in the demand for money in the last 30 years largely put an end to such proposals. Innovation in the financial world also made the definition of money much more difficult to nail down, as people could more easily switch between a greater variety of different forms of money when circumstances changed. Emerging out of those discussions was an emphasis on targeting the price level or the rate of inflation rather than monetary aggregates per se. With a stable demand for money, targeting the money supply implies a target for the price level or inflation rate. With the demand for money less stable, targeting the money supply means a price level that varies inversely to the demand for money.

As noted earlier, the Fed currently has a dual mandate that involves both unemployment and inflation. A price-level target could take the form of dropping the unemployment part of the dual mandate and simply making either the absolute level of prices or the growth rate in the price level the direct target. With a single mandate focused on the price level, the Fed would no longer try to affect the unemployment rate at the cost of higher levels of inflation. Given the strong argument that the unemployment rate is the result of real factors in the economy and not monetary policy, in all but perhaps the very short run attempting to target the unemployment rate is likely to fail and to generate inflation in the process.

Targeting the price level or inflation rate presents two challenges. First, the central bank must have sufficient confidence in its model to know how to change policy tools appropriately to generate the desired effects on the money supply and thereby the price level. The instability in money demand and the increased variety of things that can be used as money pose challenges for price-level or inflation-rate

targeting as well. Any monetary policy regime that requires the central bank to make conscious adjustments in the monetary base, as opposed to a fixed growth rate, will face similar problems.

The second concern with price-level targets is that they can be in conflict with maintaining monetary equilibrium. As noted earlier, ideal monetary policy does not avoid all changes in the price level, only those generated by monetary disequilibria. Changes in the price level caused by changes in productivity are *desirable*. If we decide the right policy is to target a stable price level, and we have rising productivity pushing prices down, then the Fed would have to react with expansionary policy. Assuming no increase in the demand for money, this reaction would lead to an excess money supply and to all the problems of inflationary monetary disequilibria, despite the Fed hitting the target of a stable price level. From a monetary disequilibrium perspective, targeting the price level or inflation rate is not the ideal goal; matching changes in the demand for money is. Even with these concerns, however, tying the Fed to a single price-level or inflation-rate mandate would be superior to the dual mandate of the status quo. Whatever the challenges of hitting a price-level target, a single mandate is much less likely to produce highly problematic degrees of monetary disequilibrium than is trying to juggle a price-level target with an unemployment-rate target that is much more difficult to reach.

The third option is to eliminate the central bank entirely and allow individual banks to produce currency in a competitive process of "free banking." The Canadian experience noted earlier, along with a similar system in Scotland and in other places at various times, all indicate that this alternative can work successfully.[25] Rather than having currency produced monopolistically by a central bank, individual banks would produce it in the same way that they currently produce their own "brands" of deposits. My checking account deposits at my bank today are actually "private money" that the bank produces competitively against the checking account dollars from your bank. Banks have developed ways of clearing those checkable liabilities through various clearinghouse arrangements. Under free banking, banks would produce and clear the currency they create through similar arrangements, just as they did historically in those systems where currency production was private. Banks developed very sophisticated institutions for coordinating and overseeing their behavior, even during times of crisis.

Perhaps the greatest advantage of a free banking system is its ability to avoid inflation and deflation. Free banks would want to make their currency and checking accounts redeemable in some sort of commodity as a way to assure customers of their value. Historically, this commodity has been gold, and I will continue under

25. On historical examples of free banking, see Kevin Dowd, ed., *The Experience of Free Banking* (London: Routledge, 1992) and Lawrence H. White, *Free Banking in Britain* (Cambridge: Cambridge University Press, 1984). On the theory of free banking, see George Selgin, *The Theory of Free Banking: Money Supply Under Competitive Note Issue* (Totowa, NJ: Rowman and Littlefield, 1988).

that assumption here. With currency and deposits redeemable in gold, customers and other banks can take any excess balances of such liabilities to the issuer for gold. Unlike the central banking scenario presented previously, there is no reason for any household or bank to want to hold balances of any *specific* currency or checking account because competition means that there are alternative banks to patronize. The result is that the redemption process works far more effectively here than under central banking, mirroring the general result that competition works better than monopoly.

Should any bank produce more money than its customers wish to hold, those customers will either bring it back to the bank directly for redemption or they will spend it, where it will most likely end up in the possession of a different bank. The other bank will not want to hold stocks of a competitor's money. Instead, it will prefer to redeem it for gold or reserves at the bank directly or at a clearinghouse, either of which will impose a cost on the competitor by taking away the gold or reserves it needs to create loans. This process of "adverse clearings" ensures that if any bank creates too much money, it will pay an economic price for it in the form of reduced reserves. Lower reserves not only limit what the bank can lend and thereby earn in interest, but insufficient reserves also put the bank at risk of not being able to pay depositors. Should a bank create too little money, it will also pay a price, but in the form of having too many reserves on hand and thereby sacrificing the interest it could earn by expanding its lending. Free banks would avoid deflation because underproducing money is costly. If we assume that banks are profit seekers, they would have every reason to avoid both inflation and deflation.

What a free banking system produces is the right degree of flexibility necessary for sound money. Because this system is separate from the government, we need not worry about political incentives working at cross-purposes with sound money. Unlike discretionary central banks, free banks do not face the lag and information problems noted earlier. With banks operating in a truly competitive market, they are able to make use of market signals, such as their reserve holdings and profits, to show them quickly and accurately whether they have produced the right quantity of money. Although banks will not get the money supply exactly right at every moment, a competitive free banking system will ensure that they know they have erred and that they have the knowledge and incentives needed to correct mistakes, and it will do so better than any alternatives. A free banking system also has the advantage of being able to respond to changes in the demand for money, while still being constrained to not over- or undercorrect, unlike the rule-bound central bank. This "flexibility within constraints" is a product of the competitive environment that free banking creates.

The one possible drawback of free banking is that it is hard to imagine political actors wanting to adopt a system that removes power from their hands and constrains their ability to deficit-spend. Without a central bank conducting open-market operations, the public must be willing to buy up bonds at a reasonable price

in order for governments to spend money they do not have. In this way, free banking not only puts constraints on monetary policy, it does the same for fiscal policy. Although these constraints are good for avoiding the bloated government and dangerous debt levels the United States is currently experiencing, they also make it less likely that the political process that benefits from the status quo will lead to the needed institutional reform.

The final option has been the subject of much discussion in the last few years: targeting either the absolute level of, or growth rate in, nominal GDP (NGDP). NGDP can be thought of as the total volume of spending unadjusted for inflation. That total volume of spending will be determined by the total quantity of money in the economy times the average number of times it changes hands. Defenders of an NGDP rule argue that fluctuations in that total volume of spending are largely responsible for fluctuations in real variables such as unemployment. They often point to the Great Depression as an example. The 30 percent decline in the money supply in the early 1930s caused NGDP to fall catastrophically, reducing the demand for labor as well as for investment and turning a potentially mild recession into the Great Depression. These defenders also point to the current slow recovery as another example, pointing out that nominal GDP growth rates are still below their pre-crisis trend and that had the Fed chosen to stick to an NGDP growth-rate rule, the crisis would not have been as deep and the recovery would not be so slow. There are a variety of methods by which NGDP could be targeted, and the debates among them are too technical for our purposes here. What matters is that NGDP targeting does have some important advantages to consider.

First, NGDP targeting is superior to price-level or inflation-rate targeting because it takes account of changes in real growth by using NGDP rather than just the price level. In that way, it is much closer to monetary equilibrium, although most defenders of NGDP targeting would want to target a growth rate in NGDP, while monetary equilibrium, strictly speaking, would mean targeting a value of NGDP. Offsetting changes in money demand with changes in the supply would lead to a constant NGDP. In either case, however, both price-level and real-growth changes are accounted for.

Second, if the central bank can successfully target NGDP (which is not clear), it will accomplish the same goal as a free banking system would. If free banking is politically challenging, then NGDP targeting might be a more feasible alternative. Again, the argument that NGDP targeting is more politically feasible than free banking presumes that NGDP targeting is workable. Critics of NGDP targeting, including many who would defend free banking, argue that the same sorts of knowledge and incentive problems that plague more discretionary forms of monetary policy would make it hard for central banks to target NGDP successfully. Can central bankers collect the requisite information (or set up the right institutions) and will they stick to that policy in the face of the temptation to deviate for political gain? Even with those concerns, it is worth asking whether NGDP targeting, though not ideal in the way

that free banking would be, is still a significant improvement over both the current policy regime and the first two alternatives presented here.

CONCLUSION

The continued failures of central banks to conduct monetary policy in a way that promotes sound money and consistent, sustainable economic growth are a good reason for the public to become more knowledgeable about monetary policy and the operation and history of monetary institutions. Even though the ideas may seem arcane and complex, their importance is immense, as money's pervasiveness in market economies magnifies dysfunctional institutions, counterproductive regulations, and policy errors. The effects of monetary mischief cannot be isolated to just one corner of the economy. Because money is half of virtually every exchange, errors in policy will spill over across the entire economy.

The widespread effects of poor monetary regulation, institutions, and policy are also why it is time to give serious consideration to more fundamental institutional changes in the monetary regime, such as those associated with free banking. Even as we recognize the challenge of political feasibility, we can also recognize the endemic problems that central banks face and the need for structural reform. The debate must get beyond merely which interest rate or money supply target is best, or what sorts of central bank tools would best achieve them. When we understand how monetary policy works and the problems central banks face in trying to get it right, we should be increasingly open to these sorts of institutional alternatives. The history of the US banking system is littered with the tragic consequences of bad regulation and poor institutions, from the panics of the 19th century, to the Great Depression, to the inflation of the 1970s and 1980s, to the housing boom and bust and ensuing recession of our own times. We cannot afford to continue these mistakes, and we must now put all the available options on the table without concern about how those who benefit from the status quo will view them. The stakes are too high and time is too short.